A peer-reviewed open-access journal
NeoBiota
Advancing research on alien species and biological invasions

NeoBiota 96: 191–210 (2024)
DOI: 10.3897/neobiota.96.129547

Research Article

Supporting early detection of biological invasions through short-term spatial forecasts of detectability

César Capinha[1,2], António T. Monteiro[1,2,3], Ana Ceia-Hasse[4,5,6]

1 *Centro de Estudos Geográficos, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa, Rua Branca Edmée Marques, 1600-276 Lisboa, Portugal*
2 *Laboratório Associado Terra, Lisboa, Portugal*
3 *Istituto di Geoscienze e Georisorse, Consiglio Nazionale delle Ricerche, Via Moruzzi 2, 56124 Pisa, Italy*
4 *CIBIO, Research Centre in Biodiversity and Genetic Resources, InBIO Associate Laboratory, School of Agronomy, University of Lisbon, 1349-017 Lisboa, Portugal*
5 *CIBIO, Research Centre in Biodiversity and Genetic Resources, InBIO Associate Laboratory, Vairão Campus, University of Porto, 4485-661 Vairão, Portugal*
6 *BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão Campus, 4485-661 Vairão, Portugal*

Corresponding author: César Capinha (cesarcapinha@edu.ulisboa.pt)

Abstract

Early detection of invasive species is crucial to prevent biological invasions. To increase the success of detection efforts, it is often essential to know when key phenological stages of invasive species are reached. This includes knowing, for example, when invasive insect species are in their adult phase, invasive plants are flowering or invasive mammals have finished their hibernation. Unfortunately, this kind of information is often unavailable or is provided at very coarse temporal and spatial resolutions. On the other hand, opportunistic records of the location and timing of observations of these stages are increasingly available from biodiversity data repositories. Here, we demonstrate how to apply these data for predicting the timing of phenological stages of invasive species. The predictions are made across Europe, at a daily temporal resolution, including in near real time and for multiple days ahead. We apply this to phenological stages of relevance for the detection of four well-known invasive species: the freshwater jellyfish, the geranium bronze butterfly, the floating primrose-willow and the garden lupine. Our approach uses machine-learning and statistical-based algorithms to identify the set of temporal environmental conditions (e.g. temperature values and trends, precipitation, snow depth and wind speed) associated with the observation of each phenological stage, while accounting for spatial and temporal biases in recording effort. Correlation between predictions from models and the actual timing of observations often exceeded values of 0.9. However, some inter-taxa variation occurred, with models using direct predictors of phenological drivers and trained on thousands of observation records outperforming those relying on indirect predictors and only a few hundred training records. The analysis of daily predictions also allowed mapping European-wide regions with similar phenological dynamics (i.e. 'phenoregions'). Our results underscore the significant potential of opportunistic biodiversity observation data in developing models capable of predicting and forecasting species phenological stages across broad spatial extents. By enhancing our current ability to anticipate the phenological stages of invasive species, this approach has the potential to significantly improve decision-making in invasion surveillance and monitoring activities.

Key words: Citizen science, early warning systems, field surveying, invasion monitoring, phenology tracking, real-time forecasting

OPEN ACCESS

Academic editor: Sven Bacher
Received: 12 June 2024
Accepted: 29 October 2024
Published: 18 November 2024

Citation: Capinha C, Monteiro AT, Ceia-Hasse A (2024) Supporting early detection of biological invasions through short-term spatial forecasts of detectability. NeoBiota 96: 191–210. https://doi.org/10.3897/neobiota.96.129547

Copyright: © César Capinha et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

Invasive alien species are a major environmental problem, severely impacting biodiversity, economies and public health (IPBES 2023). As human activities continue to transport and introduce alien organisms outside native regions (Hulme 2021; Capinha et al. 2023), the number of new invasions is expected to grow (Seebens et al. 2021), increasing the magnitude of their impacts (Bacher et al. 2023). As a result of this, successful invasion prevention efforts are key for safeguarding biodiversity, agriculture, economic development and human well-being (Vilà and Hulme 2017; IPBES 2023). In this regard, it is particularly important to detect non-native species in the early stages of the invasion process, as it significantly improves the effectiveness of control measures (Tobin et al. 2014; Larson et al. 2020; Martinez et al. 2020).

Current efforts in the surveillance and early detection of alien species encompass a large diversity of approaches, including camera and chemical traps, eDNA analysis, remote sensing and visual surveys conducted by experts and citizen scientists (Larson et al. 2020). Each of these approaches has distinct strengths and limitations and optimal outcomes are likely achieved through the integration and assimilation of their collective data (Larson et al. 2020; Fricke and Olden 2023). A key factor in improving the success and cost-effectiveness of many of these methods is understanding when invasive species are in phenological stages that maximise detection efficiency. For instance, remote sensing often targets periods of higher species conspicuity, such as plant flowering or increased greenness, to enhance detection accuracy (e.g. Müllerová et al. (2017); Wijesingha et al. (2020)). Similarly, trap-based surveillance programmes are often tailored to specific life stages of the target species and the deployment of these traps aims to coincide with the expected timing of these stages (Takeuchi et al. 2023; Nguyen et al. 2024). Likewise, visual-based field surveys – either performed by experts or citizen scientists – also greatly benefit from knowing when phenological stages of higher visual conspicuousness are reached (e.g. the adult phase of an invasive insect species or the blooming of an invasive plant), sustaining the development of surveys and monitoring activities when detectability is higher (Pocock et al. 2023).

Despite the importance of understanding the optimal timing for surveillance and early detection, information on species detectability levels is often unavailable, inadequate or of limited value. For most invasive species, including highly problematic ones, the available information on these levels typically consists of dates of relevant life cycle stages observed in other regions (e.g. EFSA et al. 2020) or the months or seasons when these stages are typically observed (Veenvliet et al. 2019). However, this type of information can overlook significant inter-regional and inter-annual differences resulting from the natural variation of drivers of phenology, such as temperature and precipitation (Godoy et al. 2009). Some exceptions exist for species for which phenological models have been developed. These models, whether process-based or data-driven, have yielded successful predictions of phenology (e.g. Barker et al. (2020); Reznik et al. (2022)), thereby supporting invasion surveillance efforts (e.g. Taylor et al. (2020); Takeuchi et al. (2023); Barker and Coop (2024)). However, applying these models to a large number of invasive species — i.e. taxonomic scalability — can present significant difficulties. Process-based models require species-specific eco-physiological parameters, which can be unavailable or hard to obtain for the species of interest (Chuine and Régnière

2017; but see Barker et al. (2020) and Takeuchi et al. (2023)). Conversely, most data-driven phenological modelling approaches rely on long-term observational data, such as phenological time series (e.g. Taylor and White (2020)), which are typically limited in spatial and taxonomic coverage (Park et al. 2021). These limitations can hinder the application of such approaches to a broad range of taxa.

Recently, we have demonstrated how temporally and spatially discrete biodiversity observation data, widely available from popular online repositories, such as the Global Biodiversity Information Facility (GBIF: https://www.gbif.org/) or iNaturalist (https://www.inaturalist.org/), can be used to estimate the timing of ecological phenomena across regions (Capinha et al. 2024). This approach is based on the concept of the phenological niche (Post 2019) and, in simple terms, involves using these data to represent the set of temporal environmental conditions under which an ecological phenomenon of interest (such as a species' phenological stage) occurs. From a practical standpoint, this can be achieved by applying statistical or machine-learning models to identify the 'envelope' of temporal environmental conditions associated with the observation of the phenomenon of interest. Once calibrated to perform this identification, the models can then be coupled with environmental predictor data (such as spatial time series of meteorological variables) and used to predict the probability of the phenomenon occurring over time (e.g. each day) and across regions.

Our previous work (Capinha et al. 2024) focused on demonstrating the conceptual and applied feasibility of this approach. Here, we aim to specifically highlight its potential for supporting efforts of early detection and monitoring of invasive species. This is done by demonstrating its use for predicting the timing of phenological stages of relevance for field surveying, at a daily resolution and for several days in advance, across Europe. We examine four well-known invasive species in Europe, offering different levels of observation data availability, and explore model accuracy, uncertainty and the regional patterns of predicted phenological dynamics. Ultimately, we demonstrate that this approach has significant potential to provide information for invasion surveillance efforts, while also offering relevant taxonomic scalability.

Methods

Collation of observation record data

We focus on four alien species that are established in Europe: the freshwater jellyfish (*Craspedacusta sowerbii*), the geranium bronze butterfly (*Cacyreus marshalli*), the floating primrose-willow (*Ludwigia peploides*) and the garden lupine (*Lupinus polyphyllus*). The levels of visual detectability for these species change considerably throughout the year. The freshwater jellyfish is presumed to be native to regions of Asia and has been introduced in most continents of the world. However, its alien distribution remains poorly known, largely because the most visible part of its life cycle involves small medusae that appear for only a few months each year (Marchessaux et al. 2021). The geranium bronze is a small butterfly native to southern Africa and currently invading parts of central and southern Europe. Like most insects, its adult (butterfly) stage has higher visibility due to increased mobility and conspicuous colours. The floating primrose-willow is an aquatic plant native to Oceania and the Americas, with invasive populations in countries of central and southern

Europe. This species produces bright yellow flowers, which facilitate its identification amongst surrounding vegetation (Booy et al. 2015). Finally, the garden lupine is a plant native to western North America that is now widespread in many temperate regions of the world, including central and northern Europe. This species produces prominent flower spikes (often violet, but sometimes also pink or white) that greatly facilitate its detection, including via remote sensing (Wijesingha et al. 2020).

Following our previously described framework (Capinha et al. 2024), we collected observation records supported by photographs for each of these species from GBIF. These records were sourced globally (i.e. not limited to Europe) to maximise the range of environmental conditions sampled. To ensure temporal consistency with the environmental predictors (see below), we used georeferenced records from a 7-year period between 2016 and 2022 that included the complete date of observation (day, month and year). GBIF is a major online repository of biodiversity observation data, including from well-known and highly participated citizen-science projects (e.g. iNaturalist.org and observation.org), which typically provide supporting visual media. As the number of records obtained from GBIF for the freshwater jellyfish was low, we also searched for observation records of this species in the USGS Non-indigenous Aquatic Species portal (https://nas.er.usgs.gov) and a few additional sources (see link for full list in the Data Resources section).

We visually checked the photographs supporting each observation record of the four species and kept only those that clearly showed the medusa stage of the freshwater jellyfish, the butterfly stage of the geranium bronze and the flowering stages of both the floating primrose-willow and the garden lupine. Records with images suggesting that the specimens were under human-care (e.g. garden lupine in places showing garden-like features) were excluded. Likewise, we also excluded GBIF records where the observation date was the first day of the month and the observation time was '00:00:00'. These are typically records where only the month and year of observation are known and the first day of the month is assigned by default, i.e. the full date of the record may not be precise (Belitz et al. 2023). In total, we obtained 240 records for freshwater jellyfish medusae, 3,879 for geranium bronze butterflies and 1,688 and 10,345 records for floating primrose-willow and garden lupine flowers, respectively (Suppl. material 1: fig. S1).

Environmental drivers

We collected a time series of global-scale maps representing daily conditions of maximum, minimum and mean temperature, accumulated precipitation, wind speed and accumulated snow. These factors are expected to be drivers of the timing of occurrence of the species life stages of interest, according to previous research (Favilli and Manganelli 2006; Ludewig et al. 2022; Marchessaux et al. 2022). The data were collected from the Global Forecast System (GFS), a large-scale NOAA weather forecast model (https://www.ncei.noaa.gov), covering the period between 15 January 2015 (the first day the data is available) to 31 December 2022. GFS data are provided at 0.25° spatial resolution, for multiple hourly intervals and for each model run that takes place at 00, 06, 12 and 18 UTC daily. The daily-resolution maps were obtained by aggregating the first six-hour values provided by each model run. This means we used the forecast for 00:00–06:00, 06:00–12:00, 12:00–18:00 and 18:00–00:00 UTC, covering the full 24 hours of each day. The temporal immediacy of this period in relation to the timing of model runs results in highly accurate weather

forecasts (NOAA 2022). Models employing these data achieve results comparable to those using climate re-analysis data - traditionally for ecological forecasting - like ERA5 (Capinha et al. 2024). Crucially, NOAA also offers real-time access to GFS data, including weather forecasts extending several days into the future, meeting our objective to deliver predictions in real-time and for short-term forecasting.

Spatial bias removal

We implemented a set of procedures to minimise potential spatial and temporal biases in the observation data. Spatial bias refers to unequal numbers of records in distinct regions, which can lead to model responses being 'dominated' by the patterns occurring in oversampled regions. Temporal observation bias arises from unequal levels of recording effort within and across years, confounding the actual temporal signal of phenological events.

To address these biases, we followed the procedures we proposed earlier (Capinha et al. 2024). Specifically, and to minimise spatial bias, we kept only one record located in the same grid cell and having the same date of observation. Next, we downsampled observations in oversampled regions. For this purpose, we used a reference grid of 250 km × 250 km squares, for which we identified squares that were upper outliers in terms of record count (i.e. n > third quartile + 1.5 * interquartile range). For these areas, we randomly subsampled a number of records equal to the outlier threshold.

Temporal bias removal

Our framework includes an optional procedure to minimise temporal bias, named 'benchmark taxa approach' (Capinha et al. 2024). This procedure uses taxa that maintain a consistent visual appearance throughout the year as indicators of variations in recorder activity (e.g. citizen scientists). We use pine species for this purpose because they are evergreen and have a fairly consistent appearance year-round. Since these taxa do not undergo significant seasonal changes, temporal variations in their recording frequency are likely due to fluctuations in recording effort rather than changes in the taxa themselves. In other words, we expect that the temporal patterns in recordings of these species serve as a proxy for general citizen-science activity levels (see further description of the rationale in Capinha et al. (2024)).

The temporal variation in the frequency of records for these taxa is related to variables expected to mediate levels of recording effort (e.g. days of the week, months of the year and weather conditions) by means of a statistical model such as a generalised linear model. Based on the relationships identified, the temporal biases in records of the phenomenon of interest can be minimised by a subsampling procedure, where records made in periods of higher levels of recording intensity receive a lower probability of being selected for model development. We demonstrated this approach previously and its application delivered similar performance to the models without using it. However, it is not clear if this outcome can be expected in the generality of phenomena. We therefore performed all the analyses using event observation data with this correction (described in Suppl. material: text S1) and without it. The results were similar in both approaches (see Results); therefore, the approach using the temporally corrected data is presented only in the Suppl. material 1.

Environmental characterisation of records

We next characterised the meteorological conditions preceding each event record. We used a total of 67 features representing multiple features of temperature (e.g. maximum, minimum and mean values, growing degree days and cold accumulation), accumulated precipitation, accumulated snow and mean wind speed for distinct preceding periods ranging from days, weeks and months up to a year (see full list in Suppl. material 1: table S1).

Additionally, we assembled a second set of data aimed at representing the meteorological conditions that are generally available in the location of each of the events (i.e. the background environmental conditions). This was performed using 'temporal pseudo-absences' (Capinha et al. 2024), which correspond to records having the same geographical coordinates as event records, but with dates drawn at random within the temporal range of the event data. A total of 12 pseudo-absence records were generated from each event record and each was characterised using the same set of 67 environmental features. The number of pseudo-absence records per observation was determined subjectively, based on preliminary assessments, which indicated a good balance between the diversity of sampled conditions and the volume of data produced.

Model training and evaluation

Prior to model fitting, we tested for the presence of multicollinearity amongst the predictors. For this purpose, we measured their variance inflation factor ('VIF') and excluded any predictor with a VIF value above 10 (Gareth et al. 2021). We then trained three machine-learning algorithms: random forests (RF), boosted regression trees (BRT) and generalised linear models with lasso regularisation (GLM-lasso), in distinguishing the conditions associated with the phenological stages of interest and those represented by the temporal pseudo-absences. These algorithms were selected because they are commonly used for ecological modelling and prediction and often rank amongst the best performing, including when transferred to new spatial settings (Zhang et al. 2019; Valavi et al. 2023).

The implementation of these models was performed in R (R Core Team 2024), using the 'randomForest' package for RF (Liaw and Wiener 2002), the 'dismo' package for BRT (Hijmans et al. 2017) and 'glmnet' for GLM-lasso (Hastie et al. 2021). We optimised several parameters within these functions to improve model fitting. Random forests used 2,000 individual trees (instead of the default 500) to increase the chances of the relatively large number of predictor variables and samples being adequately represented in the final ensemble. For the BRT models, the number of trees in each ensemble was automatically determined by the `gbm.step` function of the 'dismo' package, with a tree complexity of 3 (allowing for interactions amongst predictor variables) and a learning rate of 0.005. Additionally, for all three algorithms, we addressed the class imbalance in the data (i.e. one event observation record for every 12 pseudo-absence records), which could skew models towards over-predicting the dominant class (pseudo-absences). For GLM-lasso and BRTs, this was made through a weighting parameter ('weights' for GLM-lasso and `site.weights` for BRTs), assigning class-proportional weights to each sample. For Random Forests, the adjustment was made using the 'sampsize' parameter of the 'randomForest' function, ensuring that each individual tree used an equal number of event observation records and temporal pseudo-absences.

To evaluate the predictive performance of the models, we used a leave-one-year-out cross-validation procedure. This involved excluding the data from one year for model calibration and using it to assess the predictive ability of models trained on the remaining years. The procedure was iterated so that the data from each year served as an evaluation set. To measure the models' performance, we used the Boyce Index, initially proposed for species distribution models (Hirzel et al. 2006). The Boyce Index measures the correlation between the frequency of observation records and predicted probability values. In the context of our work, a strong positive correlation implies that event observations are concentrated in periods predicted with higher probability values, whereas correlations near zero indicate predictions akin to those expected from a model that assigns probability values randomly throughout the year. We preferred this metric over discrimination-based metrics, like the commonly-used area under the receiver operating characteristic curve (Pontius and Parmentier 2014), because these latter metrics would involve assessing the misclassification of temporal pseudo-absences. However, in the context of our work, these pseudo-absences capture the environmental variation that is available, including periods that are potentially suitable for the occurrence of our focal phenological stages. This becomes particularly critical for models of phenological stages with extended seasons of occurrence, where temporal pseudo-absences during suitable periods are more likely. Consequently, this would likely result in an underestimation of the models' true predictive performance due to the misclassification of pseudo-absences.

We performed the Boyce Index calculations using the *ecospat.boyce* function from the *ecospat* R package (Broennimann et al. 2015), employing the Pearson correlation coefficient as a metric. The measurements were performed for the three algorithms and for an ensemble of these predictions corresponding to their average.

Real-time, days-ahead forecasts of event probability

The ability to predict species' phenological stages several days in advance can guide decision-making on the optimal timing of early detection efforts for invasive species (Barker et al. 2020; Crimmins et al. 2020; Barker and Coop 2024). Consequently, an important objective of our work was to use the models developed to provide real-time phenological forecasts for several days into the future. We performed this, based on GFS weather forecasts and for up to 9 days into the future, a forecast horizon that is broad enough to guide short-term surveying decisions, but which is not too close to the limit of the GFS forecast horizon (16 days), allowing to accommodate technical challenges such as server downtime and decreased temporal resolution of GFS data after 10 days. These forecasts were implemented in real-time, providing short-term forecasts of detectability and can be accessed at https://www.natureforecast.org.

An important question is whether the forecasts lose accuracy as they extend further into the future and, if so, to what extent. To address this, we calculated the Boyce Index for phenology forecasts derived from GFS weather data produced immediately before the target day (i.e. the 18:00 UTC run of the day before). We then compared these with forecasts based on weather data generated 3, 6 and 9 days in advance. This assessment covered a 10-month period from 1 June 2023 to 31 March 2024, corresponding to the timeline from the real-time deployment of the forecasting models to the writing of this work. As evaluation data, we gathered

observation records for this period from GBIF, keeping only those that represented the life stages of interest and performing the same initial data-cleaning procedures as for calibration data (i.e. removing records without full date attributes and duplicates in space and time). Only records in Europe were considered, matching the geographical focus of the work (i.e. where the four species are invasive). Observation records for *L. polyphyllus* for this period were highly voluminous (> 19,000 records). To reduce the time resources needed to visually identify the life stage of each observation, a subset of 1000 randomly selected records was considered for processing.

Mapping phenoregions

Identifying regions with similar year-to-year phenological patterns ("phenoregions"; White et al. (2005)) allows defining areas where surveillance programmes for invasive species could be implemented in the same periods. To identify these regions, we used the phenology predictions from RF models at 5-day intervals for the 7-years period (from 2016 to 2022), across Europe. We used the predictions from this algorithm because it consistently performed well, ranking as the best or second-best across all species (see Results). To minimise the risk of extrapolation errors, i.e. making predictions for conditions far from those represented in the training data, we only made predictions for regions with Köppen climate classes (from Beck et al. (2018)) where each phenological stage had at least five observation records. Although simple, this method allows us to discriminate between regions with higher or lower propensities for model extrapolation. While this approach was effective for our purposes, future research could benefit from incorporating finer-scale assessments of extrapolation.

We applied a k-means algorithm to cluster regions based on the temporal variation in predicted values, using the 'elbow' method to determine the optimal number of clusters (Syakur et al. 2018). For each identified region, we then calculated the mean and standard deviation of inter-annual variation.

Results

Predictive performances

Overall, the predictions from models (Fig. 1, Suppl. material 1: figs S2, S3) achieved high performances. The predictions of imago-stage of the Geranium bronze attained a cross-model median correlation with the timing of observations of 0.95. GLM-Lasso achieved the highest median performance ($r = 0.97$), followed by RF ($r = 0.96$), multi-model ensemble ($r = 0.95$) and BRT ($r = 0.94$) (Fig. 2). Pairwise Kruskal-Wallis tests indicate non-significant differences in these performances (i.e. $p \geq 0.05$; Suppl. material 1: table S2). Likewise, the flowering timings of the floating primrose-willow and the garden lupine were also well captured by the models, with median cross-model correlations of 0.94 and 0.91, respectively. For the floating primrose-willow, the multi-model ensembles, RF and GLM-Lasso models had the best performance (median $r = 0.94$), with BRT performing slightly the worst (median $r = 0.88$). However, as before, these differences were not statistically significant (Kruskal-Wallis $p \geq 0.05$; Suppl. material 1: table S2). In contrast, for the garden lupine, RF displayed a statistically significant superiority (median $r = 0.96$; Kruskal-Wallis $p \leq 0.05$; Suppl. material 1: table S2), over remaining algorithms. BRT was second best (median $r = 0.92$),

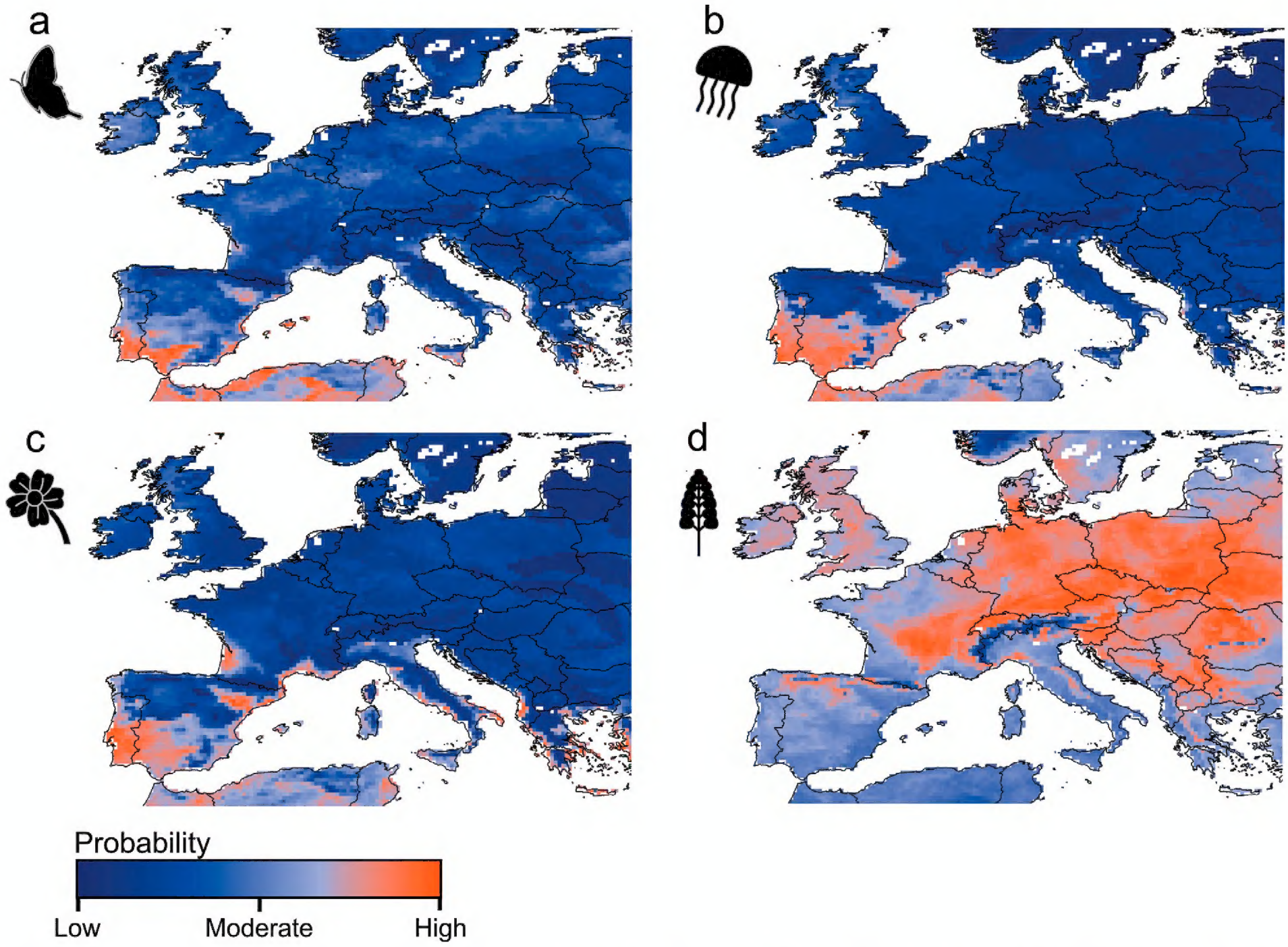

Figure 1. Examples of daily predictions obtained from the modelling approach. These represent the probability of occurrence for each modelled phenological stage for four species (imago-stage of the Geranium bronze [**a**]; medusae of the freshwater jellyfish [**b**], flowering of the floating primrose-willow [**c**] and flowering of garden lupine [**d**]), for 1 July 2023. Predictions were obtained using random forests, the best performing algorithm, trained with observational data corrected for spatial bias.

followed by multi-model ensemble and GLM-Lasso (median r for both = 0.88), but with no statistically significant differences amongst these (Kruskal-Wallis $p \geq 0.05$; Suppl. material 1: table S2). Predictive performance for the medusae stage of the freshwater jellyfish was the lowest, but still reasonable – with a median correlation of 0.73 across models (Fig. 2). Predictions from the ensemble approach demonstrated a substantially higher performance (median r = 0.82), followed by RF (median r = 0.79), GLM-Lasso (r = 0.64) and BRT (r = 0.42), but these differences were also not statistically significant (Suppl. material 1: table S2).

Relevantly, models trained with data addressing both spatial and temporal biases showed similar predictive performances as those addressing only spatial biases (Suppl. material 1: fig. S4). Specifically, the predictions of imago-stage of the Geranium bronze retained the highest cross-model median correlation with the timing of observations (0.95), followed by the garden lupine (0.91), the floating primrose-willow (0.90) and the freshwater jellyfish (0.67). Performances for the Geranium bronze were highest for GLM-Lasso, RF and multi-model ensemble (median r = 0.95), followed by BRT (median r = 0.90). For the floating primrose-willow, the best predictions were achieved by multi-model ensembles and RF (median r = 0.91), followed by GLM-lasso (median r = 0.90) and BRT (median r = 0.87). For the garden lupine, the best performances were achieved by RF

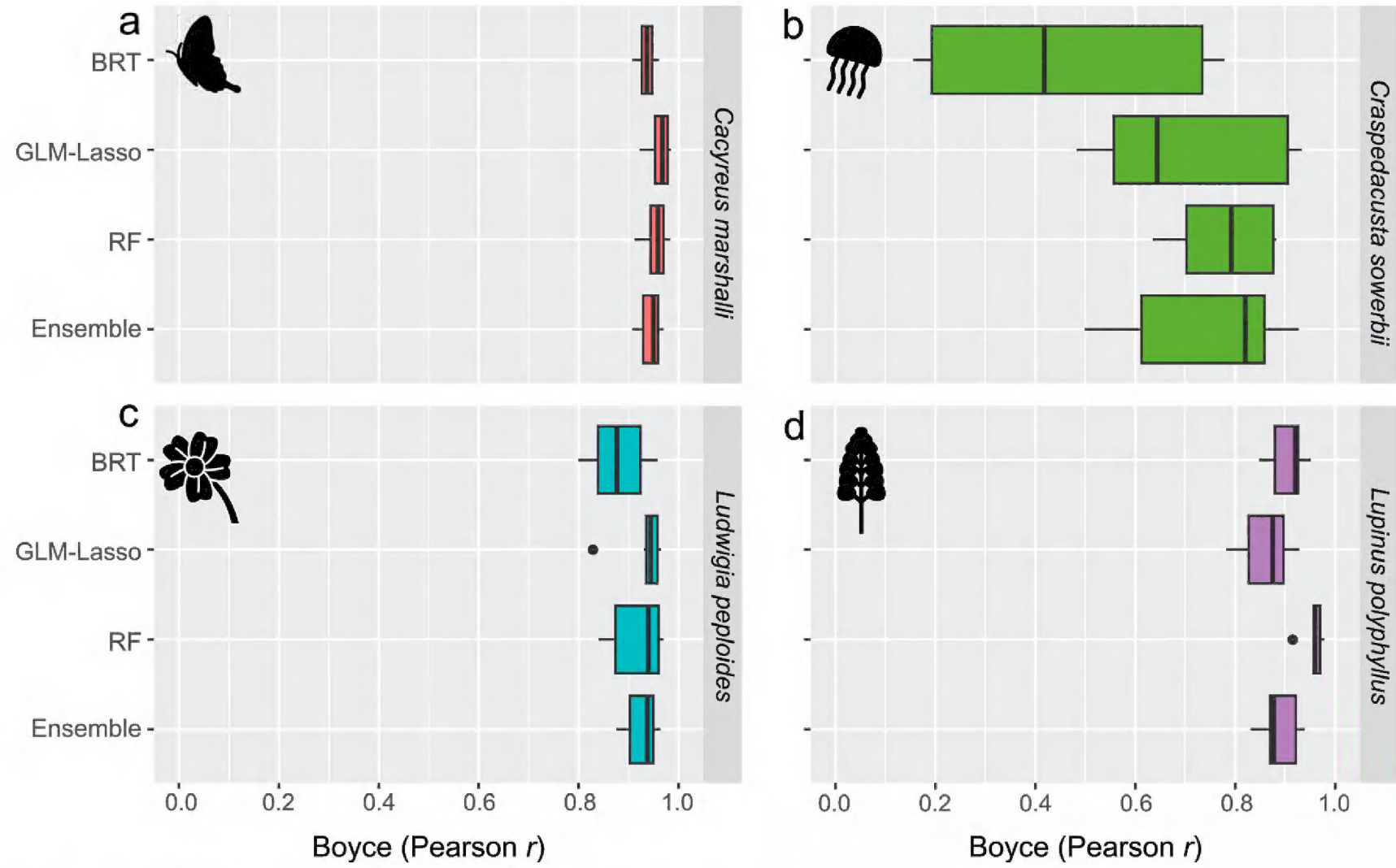

Figure 2. Results of Boyce Index corresponding to Pearson correlation values between predicted probabilities of event occurrence and the frequency of event observation records of the imago stage of the Geranium bronze butterfly (**a**), medusae of the freshwater jellyfish (**b**), the flowering of floating primrose-willow (**c**) and the garden lupine (**d**). The boxplots represent the variation of correlation values assessed for 7 years (2016 to 2022), using three modelling algorithms (boosted regression trees, BRT; generalised linear models with lasso regularisation, GLM-Lasso; random forest, RF) and an ensemble of previous algorithms (Ensemble), trained with observation data corrected for spatial bias.

(median $r = 0.97$), followed by multi-model ensembles and BRT (median $r = 0.90$) and by GLM-Lasso (median $r = 0.88$). Finally, for the freshwater jellyfish, the best performances were achieved by RF and GLM-Lasso (median $r = 0.68$), followed by multi-model ensembles (median $r = 0.66$) and BRT (median $r = 0.44$). As for models using observation data corrected for spatial bias only, the only statistically significant difference amongst algorithms was verified for the garden lupine, with RF significantly outperforming remaining ones (Suppl. material 1: table S2).

We also assessed the performance of days-ahead forecasts across Europe over a 10-month period (Fig. 3; Suppl. material 1: fig. S5). For the freshwater jellyfish and the floating primrose-willow, the observation records used for model evaluation during this period were relatively limited (25 and 42, respectively), while they were more abundant for the Geranium bronze (1,187) and the garden lupine (126 records from a random sample of 1,000 processed records). Overall, the forecasts demonstrated a high correlation with the timing of observed life stages modelled, with values generally close to 0.8 or higher (Fig. 3; Suppl. material 1: fig. S5). Additionally, there was no apparent pattern of decreased model performance as the forecast horizon extended (e.g. from 1 day ahead to 9 days ahead).

Regions of similar phenology ('phenoregions') and associated temporal patterns.

Using the predictions of the random forest algorithm trained with observation data corrected for spatial bias, we identified five phenoregions for the Geranium bronze butterfly in Europe (Fig. 4a), with the peak occurrence ranging from early June to late

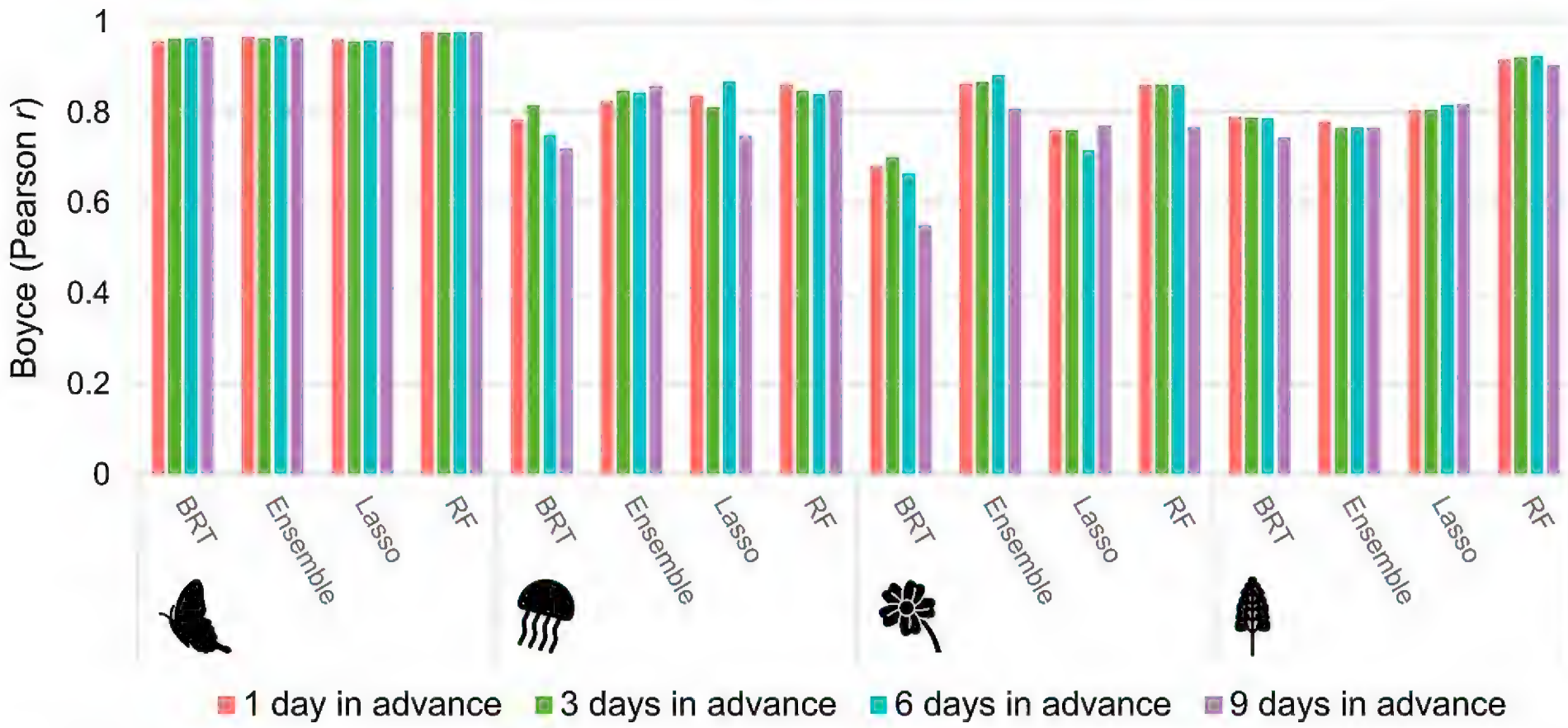

Figure 3. Boyce Index values for forecasts made 1, 3, 6 and 9 days in advance. These values correspond to the Pearson correlation coefficient between predicted probabilities of event occurrence from July 2023 to March 2024 and the frequency of event observations recorded during the same period. The values are reported for three modelling algorithms—boosted regression trees (BRT), generalised linear models with lasso regularisation (Lasso) and random forest (RF)—as well as an ensemble of these algorithms (Ensemble), all trained with observation data corrected for spatial bias.

November in the southernmost region (Fig. 4b). The remaining regions have peak occurrences between August and September, with northern latitudes having a shorter season of occurrence and the opposite being true for southern latitudes. Recorded observations of butterflies of this species are concentrated in the summer (Fig. 4c), especially in the mid- to higher latitude regions, in agreement with the predictions.

Predictions of the timing of occurrence of medusae of the freshwater jellyfish were clustered into four phenoregions, peaking between late August and September. However, southern regions exhibit higher probabilities of occurrence over substantially longer periods and, conversely, shorter periods are predicted for northern regions. Medusae observations take place in the months of predicted peaks, except for the mid-latitude region covering most of Central Europe, where they concentrate in September and November — a period when the predicted probabilities are already declining.

For the floating primrose-willow, four regions were identified (Fig. 4g), with the southern regions showing earlier and more extended flowering periods (Fig. 4h). The timing of observational data matches well with the predictions for the region including most of Mediterranean Europe, although very few records are available for the other regions (Fig. 4i).

Predicted timings of flowering for the garden lupine were classified into four phenoregions (Fig. 4j), with probabilities peaking from early May to late June and occurring earlier in southern regions (Fig. 4k). Observational data also largely agree with the predicted patterns, especially for the two northernmost regions, which have most records (Fig. 4l).

Importantly, identified phenoregions and associated temporal patterns were strikingly similar to those obtained, based on models calibrated with temporally unbiased observation data (Suppl. material 1: fig. S6). This further suggests the robustness of predictions to temporally varying levels of recording effort.

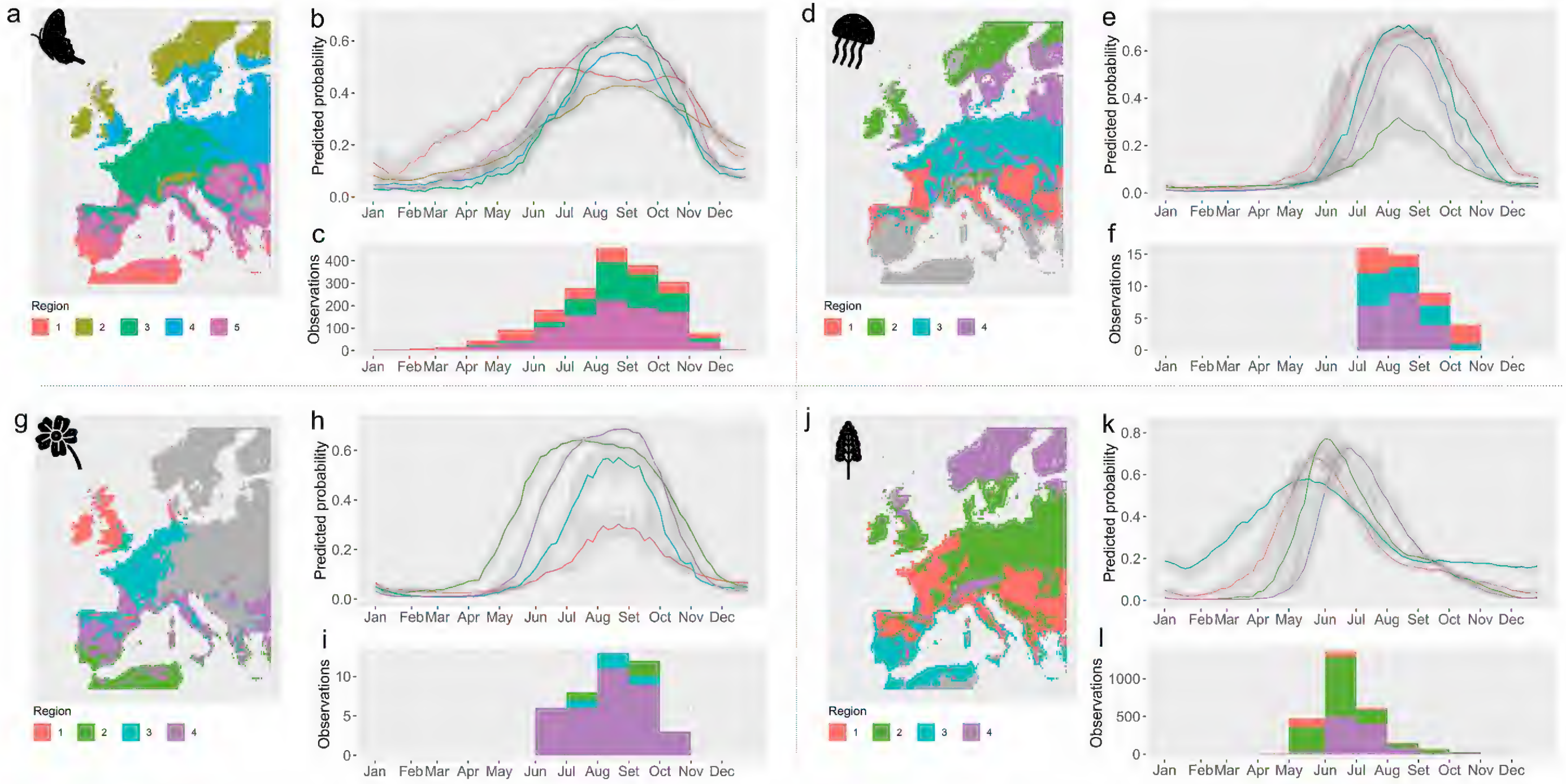

Figure 4. Regional patterns of predicted and observed timings of the emergence of the imago stage in the Geranium bronze butterfly (panels **a–c**), the occurrence of medusae in the freshwater jellyfish (panels **d–f**) and the flowering phases of floating primrose-willow (panels **g–i**) and garden lupine (panels **j–l**). The maps display areas having similar phenological dynamics ('phenoregions'), based on daily projections at 5-days intervals from 2016 to 2022. Grey areas represent Köppen climate classes where fewer than five observations of the species were made. These classes were not included in the analysis to minimise the risk of model extrapolation. Time series depict the inter-annual mean probabilities of occurrence of each event, along with their ranges (grey shading), for each region throughout the year. Histograms show the monthly frequency of effectively observed occurrences of each life stage within each phenoregion. Predictions and phenoregion delineations were made also for areas where the species have not yet been recorded, leading to observation records being absent from the histograms for certain regions.

Discussion

Efforts for the early detection of biological invasions greatly benefit from understanding when and where invasive species enter life cycle stages that enhance detectability (Müllerová et al. 2017; Barker and Coop 2024). Here, we have demonstrated that temporally discrete biodiversity observation data, as made widely available in well-known public repositories such as GBIF and contributed by citizen science (Bonney 2021; Heberling et al. 2021), can be used to estimate the timing of occurrence of such stages, including across broad spatial regions (e.g. continents) and at very high temporal resolution (e.g. daily).

Our approach was demonstrated by modelling a distinct life stage for each of four species. The results obtained demonstrate a varying ability of trained models to identify the temporal environmental variation associated with observing the life stages of interest. For three of the life stages modelled, the agreement between predictions and the timing of observation was, across models, very high, with median correlations regularly above 0.9. However, for one of them (the medusae stage of the freshwater jellyfish), the performance was lower (though still high; cross-model correlations = 0.73 and 0.67). This lower performance could be partly attributed to the use of terrestrial predictors (e.g. air temperature, wind speed and snow cover), which serve as indirect proxies for aquatic conditions and limit the ability of models to capture phenological drivers with high precision. For aquatic species, variables such as water temperature and resources are likely critical drivers of phenology (Thackeray et al. 2016; Woods et al. 2022), but were not included in our models due to a lack of data meeting the spatial extent and temporal resolution required. Therefore, while our approach demonstrates that good predictive performances can be achieved, incorporating predictors that accurately represent the primary drivers of phenology will likely be crucial.

The lower performance observed for medusae of the freshwater jellyfish also coincides with the lowest number of observation records available (few hundreds), significantly fewer than those available for the remaining life stages (in the order of thousands). This also suggests that, as with the generality of modelled phenomena, the size of the training data can be a limiting feature. Indeed, the predictions allowed by our approach, which can be made daily and over wide geographical areas, involve a high dimensionality of conditions in the prediction space resulting from the multiple states of preceding conditions for each environmental variable and their joint combination. Hence, it is a reasonable expectation that the calibration data should be necessarily large in number to represent this variability; otherwise, extrapolation may occur and the uncertainty of predictions will be higher and possibly also less accurate (Yates et al. 2018).

In this work, we did not explicitly quantify extrapolation, as it presents significant challenges in models that deal simultaneously with spatial and temporal variation. Properly assessing extrapolation in this context requires considering both its magnitude — ideally weighted by the relative importance of each predictor in the models — and its temporal recurrence. However, commonly used methods like Mahalanobis distance (Mesgaran et al. 2014) and mobility orientated-parity (Owens et al. 2013) do not adequately address these complexities. Due to these limitations, we chose not to perform a detailed extrapolation analysis, acknowledging the need for more advanced techniques in future research. Despite this, it is very positive to verify that sample sizes on the order of magnitude of a few thousands (as available for the remaining species), consistently deliver very good predictions. Indeed, the number of

invasive species for which life stages of relevance for early detection and monitoring are represented by similar amounts of records can be presumed to easily reach several thousands, considering the enormous volumes of biodiversity data becoming available each year, particularly through citizen science (Bonney 2021; Heberling et al. 2021). This highlights the potential taxonomic scalability of our approach.

It is well-known that multi-sourced opportunistic biodiversity observation data, as used to train our models, can suffer from substantial spatial and temporal bias, often hindering efforts of use for prediction (Isaac and Pocock 2015; Brown and Williams 2019). We applied a set of procedures to minimise the effect of these biases. Specifically, we corrected for the temporal bias component (see also Capinha et al. (2024)). However, we also found that models having this correction and those without it are largely equivalent in terms of predictions. This was also verified in our previous work demonstrating the general modelling workflow (Capinha et al. 2024), where we interpreted these results by virtue of the way models are trained. Conceptually, our models contrast two point clouds in the environmental space, one representing the conditions associated with the occurrence of the life stage of interest and the other the generality of conditions available in the regions of its occurrence. As random records ('pseudo-absences') are generated at a fixed ratio to observation records, observations from periods with lower recording effort will have a similar relative influence as those from periods with higher recording effort. This does not mean that predictions made for the least represented conditions will be as accurate as those made for more populated spaces of the environmental space. Rather, it means that having more populated spaces of the environmental space will not direct the models' inference towards these conditions.

Our approach is also capable of forecasting the probability of occurrence of the life stages assessed, in real time and for several days into the future. This capacity is perhaps the most impactful aspect of our work for practical applications. Providing these forecasts daily and across extensive areas (such as the European continent in this case) could support a variety of decision-making processes related to the timing and efficacy of invasive species detection efforts. Of relevance, we also observed that the performance of forecasts remains largely stable over the forecast horizon considered. This consistency may result from the relatively brief time span considered (nine days) and the inherent temporal correlation amongst phenological events, which tend to unfold gradually and slowly over time (considering the daily temporal resolution used).

The identification of phenoregions (i.e. regions sharing similar phenological dynamics), as allowed through the spatial clustering of predictions from our framework, can also be of great interest to support invasion surveillance and decision-making. Environmental managers are often left wondering which time of the year specific life stages of species will occur. Most of the information available (when available) is found in technical and scientific literature and typically indicates the months or seasons of occurrence at broad geographical resolutions, for example, a country, group of countries or a continent. For example, a highly comprehensive recent work on invasive species in the forests of Europe (Veenvliet et al. 2019) highlights the months of expected occurrence of phenological stages with the highest visibility for each species across Europe as a whole. This information is much welcomed by managers and is often the best available for most species, but it well illustrates the difficulties of obtaining high temporal and spatial resolution data on the timing of phenological stages. The maps obtained with our approach go a step further, identifying these regions at the level of individual grid cells, with temporal resolutions as fine as daily and by accounting for inter-annual variability.

In conclusion, our work demonstrates the potential of widely available, temporally discrete biodiversity observation data for estimating the timing of life stages relevant to invasive species detectability. With the increasing volume of media-supported biodiversity observation data being published, the number and diversity of invasive species for which these estimates can be produced are substantial. Furthermore, these estimates can be delivered at high spatial resolutions across wide areas, in real time and for several days into the future, providing timely decision support for numerous managers tasked with planning surveillance and early detection measures. Increasing the number of invasive species covered, while continuously refining these estimates, will likely contribute significantly to global efforts in the proactive prevention of biological invasions.

Acknowledgements

We thank Dr. Brittany Barker and an anonymous reviewer for their valuable suggestions, which helped to improve this work.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

C.C. was supported by Portuguese National Funds through Fundação para a Ciência e a Tecnologia through support to CEG/IGOT Research Unit (UIDB/00295/2020 and UIDP/00295/2020) and by the EuropaBON project, funded by European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 101003553.

Author contributions

CC: Conceptualisation, Methodology, Validation, Formal analysis, Resources, Data Curation, Writing - Original draft, Project administration, Funding Acquisition. AC and AM: Writing - Review and Editing.

Author ORCIDs

César Capinha https://orcid.org/0000-0002-0666-9755

Data availability

Spatial weather data used for this work are publicly available online from NSF NCAR Research Data Archive (RDA) n (https://rda.ucar.edu/datasets/d084001/). Event observation data are available from USGS Non-indigenous Aquatic Species (https://nas.er.usgs.gov/queries/factsheet.aspx?SpeciesID=1068) and GBIF with DOIs: https://doi.org/10.15468/dl.3fve6q, https://doi.org/10.15468/dl.9drr85, https://doi.org/10.15468/dl.h5amhh, https://doi.org/10.15468/dl.krbguc, https://doi.org/10.15468/dl.2kbnxv, https://doi.org/10.15468/dl.ycjsyz, https://doi.org/10.15468/dl.7q6rke, https://doi.org/10.15468/dl.b9sze9 and https://doi.org/10.15468/dl.uh8apz. Additional data sources and R code are publicly available on Zenodo (https://doi.org/10.5281/zenodo.13847953).

References

Bacher S, Galil B, Nunez M, Ansong M, Cassey P, Dehnen-Schmutz K, Fayvush G, Hiremath A, Ikegami M, Martinou A (2023) Impacts of invasive alien species on nature, nature's contributions to people, and good quality of life. In: Roy HE, et al. (Eds) IPBES Invasive Alien Species Assessment, 391–559.

Barker BS, Coop L (2024) Phenological Mapping of Invasive Insects: Decision Support for Surveillance and Management. Insects 15(1): 6. https://doi.org/10.3390/insects15010006

Barker BS, Coop L, Wepprich T, Grevstad F, Cook G (2020) DDRP: Real-time phenology and climatic suitability modeling of invasive insects. PLoS ONE 15(12): e0244005. https://doi.org/10.1371/journal.pone.0244005

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5(1): 1–12. https://doi.org/10.1038/sdata.2018.214

Belitz MW, Larsen EA, Shirey V, Li D, Guralnick RP (2023) Phenological research based on natural history collections: Practical guidelines and a lepidopteran case study. Functional Ecology 37(2): 234–247. https://doi.org/10.1111/1365-2435.14173

Bonney R (2021) Expanding the impact of citizen science. Bioscience 71(5): 448–451. https://doi.org/10.1093/biosci/biab041

Booy O, Wade M, Roy H (2015) Field guide to invasive plants and animals in Britain. Bloomsbury Publishing.

Broennimann O, Di Cola V, Guisan A (2015) ecospat: Spatial ecology miscellaneous methods. R package version 1. https://doi.org/10.32614/CRAN.package.ecospat

Brown ED, Williams BK (2019) The potential for citizen science to produce reliable and useful information in ecology. Conservation Biology 33(3): 561–569. https://doi.org/10.1111/cobi.13223

Capinha C, Essl F, Porto M, Seebens H (2023) The worldwide networks of spread of recorded alien species. Proceedings of the National Academy of Sciences of the United States of America 120(1): e2201911120. https://doi.org/10.1073/pnas.2201911120

Capinha C, Ceia-Hasse A, de-Miguel S, Vila-Viçosa C, Porto M, Jarić I, Tiago P, Fernández N, Valdez J, Ian M, Henrique Miguel P (2024) Using citizen science data for predicting the timing of ecological phenomena across regions. Bioscience 74(6): 383–392. https://doi.org/10.1093/biosci/biae041 [Preprint available from: https://doi.org/10.1101/2023.05.05.539567]

Chuine I, Régnière J (2017) Process-based models of phenology for plants and animals. Annual Review of Ecology, Evolution, and Systematics 48(1): 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706

Crimmins TM, Gerst KL, Huerta DG, Marsh RL, Posthumus EE, Rosemartin AH, Switzer J, Weltzin JF, Coop L, Dietschler N, Herms DA, Limbu S, Trotter RT III, Whitmore M (2020) Short-term forecasts of insect phenology inform pest management. Annals of the Entomological Society of America 113(2): 139–148. https://doi.org/10.1093/aesa/saz026

EFSA, Lázaro E, Parnell S, Civera AV, Schans J, Schenk M, Abrahantes JC, Zancanaro G, Vos S (2020) General guidelines for statistically sound and risk-based surveys of plant pests. Wiley Online Library.

Favilli L, Manganelli G (2006) Life history of *Cacyreus marshalli*, a South African species recently introduced into Italy. Bollettino della Società Entomologica Italiana 138: 51–61.

Fricke RM, Olden JD (2023) Technological innovations enhance invasive species management in the Anthropocene. Bioscience 73(4): 261–279. https://doi.org/10.1093/biosci/biad018

Gareth J, Witten D, Hastie T, Tibshirani R (2021) An introduction to statistical learning: with applications in R. 2nd edn. Spinger. https://link.springer.com/book/10.1007/978-1-0716-1418-1

Godoy O, Castro-Díez P, Valladares F, Costa-Tenorio M (2009) Different flowering phenology of alien invasive species in Spain: Evidence for the use of an empty temporal niche? Plant Biology 11(6): 803–811. https://doi.org/10.1111/j.1438-8677.2008.00185.x

Hastie T, Qian J, Tay K (2021) An introduction to glmnet. CRAN R Repository 5: 1–35. https://cran.r-project.org/package=glmnet

Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D (2021) Data integration enables global biodiversity synthesis. Proceedings of the National Academy of Sciences of the United States of America 118(6): e2018093118. https://doi.org/10.1073/pnas.2018093118

Hijmans RJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ (2017) Package 'dismo.' Circles 9: 1–68. https://rspatial.org/raster/sdm/

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199(2): 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hulme PE (2021) Unwelcome exchange: International trade as a direct and indirect driver of biological invasions worldwide. One Earth 4(5): 666–679. https://doi.org/10.1016/j.oneear.2021.04.015

IPBES (2023) Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, Germany.

Isaac NJ, Pocock MJ (2015) Bias and information in biological records. Biological Journal of the Linnean Society. Linnean Society of London 115(3): 522–531. https://doi.org/10.1111/bij.12532

Larson ER, Graham BM, Achury R, Coon JJ, Daniels MK, Gambrell DK, Jonasen KL, King GD, LaRacuente N, Perrin-Stowe TI, Reed EM, Rice CJ, Ruzi SA, Thairu MW, Wilson JC, Suarez AV (2020) From eDNA to citizen science: Emerging tools for the early detection of invasive species. Frontiers in Ecology and the Environment 18(4): 194–202. https://doi.org/10.1002/fee.2162

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2: 18–22. https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf

Ludewig K, Klinger YP, Donath TW, Bärmann L, Eichberg C, Thomsen JG, Görzen E, Hansen W, Hasselquist EM, Helminger T, Kaiskog F, Karlsson E, Kirchner T, Knudsen C, Lenzewski N, Lindmo S, Milberg P, Pruchniewicz D, Richter E, Sandner TM, Sarneel JM, Schmiede R, Schneider S, Schwarz K, Tjäder Å, Tokarska-Guzik B, Walczak C, Weber O, Żołnierz L, Eckstein RL (2022) Phenology and morphology of the invasive legume *Lupinus polyphyllus* along a latitudinal gradient in Europe. NeoBiota 78: 185–206. https://doi.org/10.3897/neobiota.78.89673

Marchessaux G, Lüskow F, Sarà G, Pakhomov EA (2021) Predicting the current and future global distribution of the invasive freshwater hydrozoan *Craspedacusta sowerbii.* Scientific Reports 11(1): 23099. https://doi.org/10.1038/s41598-021-02525-3

Marchessaux G, Lüskow F, Bejean M, Pakhomov EA (2022) Increasing temperature facilitates polyp spreading and medusa appearance of the invasive hydrozoan *Craspedacusta sowerbii.* Biology (Basel) 11(8): 1100. https://doi.org/10.3390/biology11081100

Martinez B, Reaser JK, Dehgan A, Zamft B, Baisch D, McCormick C, Giordano AJ, Aicher R, Selbe S (2020) Technology innovation: Advancing capacities for the early detection of and rapid response to invasive species. Biological Invasions 22(1): 75–100. https://doi.org/10.1007/s10530-019-02146-y

Mesgaran MB, Cousens RD, Webber BL (2014) Here be dragons: A tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. Diversity & Distributions 20(10): 1147–1159. https://doi.org/10.1111/ddi.12209

Müllerová J, Brůna J, Bartaloš T, Dvořák P, Vítková M, Pyšek P (2017) Timing Is Important: Unmanned Aircraft vs. Satellite Imagery in Plant Invasion Monitoring. Frontiers in Plant Science 8: 887. https://doi.org/10.3389/fpls.2017.00887

Nguyen H, Chu L, Liebhold AM, Epanchin-Niell R, Kean JM, Kompas T, Robinson AP, Brockerhoff EG, Moore JL (2024) Optimal allocation of resources among general and species-specific tools for plant pest biosecurity surveillance. Ecological Applications 34(3): e2955. https://doi.org/10.1002/eap.2955

NOAA (2022) WPC's Model Diagnostics and Verification Page. https://www.wpc.ncep.noaa.gov/html/model2.shtml [June 3, 2024]

Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega A, Hensz CM, Myers CE, Peterson AT (2013) Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. Ecological Modelling 263: 10–18. https://doi.org/10.1016/j.ecolmodel.2013.04.011

Park DS, Newman EA, Breckheimer IK (2021) Scale gaps in landscape phenology: Challenges and opportunities. Trends in Ecology & Evolution 36(8): 709–721. https://doi.org/10.1016/j.tree.2021.04.008

Pocock MJ, Adriaens T, Bertolino S, Eschen R, Essl F, Hulme PE, Jeschke JM, Roy HE, Teixeira H, De Groot M (2023) Citizen science is a vital partnership for invasive alien species management and research. iScience. https://doi.org/10.1016/j.isci.2023.108623

Pontius Jr RG, Parmentier B (2014) Recommendations for using the relative operating characteristic (ROC). Landscape Ecology 29(3): 367–382. https://doi.org/10.1007/s10980-013-9984-8

Post E (2019) Time in Ecology: A Theoretical Framework [MPB 61]. Princeton University Press. https://doi.org/10.2307/j.ctv3s8sns

R Core Team (2024) R: A language and environment for statistical computing. https://www.r-project.org/

Reznik SY, Karpun NN, Zakharchenko VY, Shoshina YI, Dolgovskaya MY, Saulich AK, Musolin DL (2022) To every thing there is a season: Phenology and photoperiodic control of seasonal development in the invasive Caucasian population of the brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Heteroptera: Pentatomidae). Insects 13(7): 580. https://doi.org/10.3390/insects13070580

Seebens H, Bacher S, Blackburn TM, Capinha C, Dawson W, Dullinger S, Genovesi P, Hulme PE, Van Kleunen M, Kühn I, Jeschke JM, Lenzner B, Liebhold AM, Pattison Z, Pergl J, Pyšek P, Winter M, Essl F (2021) Projecting the continental accumulation of alien species through to 2050. Global Change Biology 27(5): 970–982. https://doi.org/10.1111/gcb.15333

Syakur M, Khotimah BK, Rochman E, Satoto BD (2018) Integration k-means clustering method and elbow method for identification of the best customer profile cluster. IOP Publishing, 012017. https://doi.org/10.1088/1757-899X/336/1/012017

Takeuchi Y, Tripodi A, Montgomery K (2023) SAFARIS: A spatial analytic framework for pest forecast systems. Frontiers in Insect Science 3: 1198355. https://doi.org/10.3389/finsc.2023.1198355

Taylor SD, White EP (2020) Automated data-intensive forecasting of plant phenology throughout the United States. Ecological Applications 30(1): e02025. https://doi.org/10.1002/eap.2025

Taylor RV, Holthuijzen W, Humphrey A, Posthumus E (2020) Using phenology data to improve control of invasive plant species: A case study on Midway Atoll NWR. Ecological Solutions and Evidence 1(1): e12007. https://doi.org/10.1002/2688-8319.12007

Thackeray SJ, Henrys PA, Hemming D, Bell JR, Botham MS, Burthe S, Helaouet P, Johns DG, Jones ID, Leech DI, Mackay EB, Massimino D, Atkinson S, Bacon PJ, Brereton TM, Carvalho L, Clutton-Brock TH, Duck C, Edwards M, Elliott JM, Hall SJG, Harrington R, Pearce-Higgins JW, Høye TT, Kruuk LEB, Pemberton JM, Sparks TH, Thompson PM, White I, Winfield IJ, Wanless S (2016) Phenological sensitivity to climate across taxa and trophic levels. Nature 535(7611): 241–245. https://doi.org/10.1038/nature18608

Tobin PC, Kean JM, Suckling DM, McCullough DG, Herms DA, Stringer LD (2014) Determinants of successful arthropod eradication programs. Biological Invasions 16(2): 401–414. https://doi.org/10.1007/s10530-013-0529-5

Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2023) Flexible species distribution modelling methods perform well on spatially separated testing data. Global Ecology and Biogeography 32(3): 369–383. https://doi.org/10.1111/geb.13639

Veenvliet JK, Veenvliet P, de Groot M, Kutnar L (2019) A field guide to invasive alien species in European forests. Institute Symbiosis, so. e. and The Silva Slovenica Publishing Centre, Slovenian Forestry Institute, Ljubljana, Slovenia.

Vilà M, Hulme PE (2017) Non-native species, ecosystem services, and human well-being. In: Vilà M, Hulme P (Eds) Impact of Biological Invasions on Ecosystem Services. Invading Nature - Springer Series in Invasion Ecology, vol 12. Springer, Cham, 354 pp. https://doi.org/10.1007/978-3-319-45121-3

White MA, Hoffman F, Hargrove WW, Nemani RR (2005) A global framework for monitoring phenological responses to climate change. Geophysical Research Letters 32(4): 2004GL021961. https://doi.org/10.1029/2004GL021961

Wijesingha J, Astor T, Schulze-Brüninghoff D, Wachendorf M (2020) Mapping invasive *Lupinus polyphyllus* Lindl. in semi-natural grasslands using object-based image analysis of UAV-borne images. PFG-Journal of Photogrammetry. Journal of Photogrammetry, Remote Sensing and Geoinformation Science 88(5): 391–406. https://doi.org/10.1007/s41064-020-00121-0

Woods T, Kaz A, Giam X (2022) Phenology in freshwaters: A review and recommendations for future research. Ecography 2022(6): e05564. https://doi.org/10.1111/ecog.05564

Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ, Ban S, Barbosa AM, Dormann CF, Elith J, Embling CB, Ervin GN, Fisher R, Gould S, Graf RF, Gregr EJ, Halpin PN, Heikkinen RK, Heinänen S, Jones AR, Krishnakumar PK, Lauria V, Lozano-Montes H, Mannocci L, Mellin C, Mesgaran MB, Moreno-Amat E, Mormede S, Novaczek E, Oppel S, Crespo GO, Peterson AT, Rapacciuolo G, Roberts JJ, Ross RE, Scales KL, Schoeman D, Snelgrove P, Sundblad G, Thuiller W, Torres LG, Verbruggen H, Wang L, Wenger S, Whittingham MJ, Zharikov Y, Zurell D, Sequeira AMM (2018) Outstanding Challenges in the Transferability of Ecological Models. Trends in Ecology & Evolution 33(10): 790–802. https://doi.org/10.1016/j.tree.2018.08.001

Zhang Z, Capinha C, Weterings R, McLay CL, Xi D, Lü H, Yu L (2019) Ensemble forecasting of the global potential distribution of the invasive Chinese mitten crab, *Eriocheir sinensis*. Hydrobiologia 826(1): 367–377. https://doi.org/10.1007/s10750-018-3749-y

Supplementary material 1

Additional information

Authors: César Capinha, António T. Monteiro, Ana Ceia-Hasse

Data type: docx

Explanation note: **text 1.** Rationale and procedures used to address temporal recording bias. **table S1.** List of the 67 features used to characterize temporal environmental conditions for observation and temporal pseudo-absence records. **table S2.** Results from pairwise Kruskal-Wallis tests assessing significant differences in the performance of predictions from distinct algorithms modelling a distinct life stage for each of four species. **fig. S1.** Location of collected records of observation of imago-stage of the Geranium bronze (*Cacyreus marshalli*) (a); medusae of the freshwater jellyfish (*Craspedacusta sowerbii*) (b), flowering of the floating primrose-willow (*Ludwigia peploides*) (c) and flowering of garden lupine (*Lupinus polyphyllus*) (d), between 2016 and 2022. **fig. S2.** Examples of daily predictions obtained from models trained with observational data corrected for spatial bias. **fig. S3.** Examples of daily predictions obtained from models trained with observational data corrected for spatial and temporal bias. **fig. S4.** Boyce index values, corresponding to Pearson correlation values between predicted probabilities of event occurrence and the frequency of event observation records for models calibrated with data corrected for spatial and temporal bias. **fig. S5.** Boyce Index values for forecasts made 1, 3, 6, and 9 days in advance from models calibrated with observation data corrected for spatial and temporal bias. **fig. S6.** Regional patterns of predicted and observed timings of the phenological stages.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.96.129547.suppl1